

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>5</sup>:</b> <b>C12Q 1/68, C12P 19/34</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 94/19492</b> <b>(43) International Publication Date:</b> 1 September 1994 (01.09.94)
<b>(21) International Application Number:</b> PCT/US94/01761 <b>(22) International Filing Date:</b> 23 February 1994 (23.02.94)  <b>(30) Priority Data:</b> 08/021,960      24 February 1993 (24.02.93)      US  <b>(71) Applicant (for all designated States except US):</b> MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH [US/US]; 200 First Street, S.W., Rochester, MN 55905 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> THIBODEAU, Stephen, N. [US/US]; 1610-21st Avenue, N.E., Rochester, MN 55906 (US). BREN, Gary, D. [US/US]; Rural Route 1, Chatfield, MN 55923 (US).  <b>(74) Agent:</b> ELLINGER, Mark, S.; Patterson & Keough, 1200 Rand Tower, 527 Marquette Avenue South, Minneapolis, MN 55402 (US).		<b>(81) Designated States:</b> CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> TUMOR-SPECIFIC GENOMIC INSTABILITY AS A PROGNOSTIC INDICATOR  <b>(57) Abstract</b>  A method for assessing the level of genomic instability in a malignant tumour is described. The method involves comparison of the structure of one or more interspersed repetitive sequence loci in tumor and corresponding normal tissue samples. Microsatellite repetitive sequences are one useful form of interspersed repetitive sequence that may be employed to characterize the genomic instability associated with some malignant tumors. This assessment provides an additional prognostic indicator for patients having such malignant tumors.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

5

10

## TUMOR-SPECIFIC GENOMIC INSTABILITY AS A PROGNOSTIC INDICATOR

This invention was made with United States of America government support under Grant No. CA 15083-18E8.1 awarded by the National  
15 Institutes of Health. The government has certain rights in the invention.

### Field of the Invention

This invention relates generally to development of clinical prognoses for cancer patients by analysis of tumor nucleic acids for evidence of  
20 genomic instability. More specifically, this invention relates to a method for comparing the structural integrity of interspersed repetitive sequences in tumor and corresponding normal tissue as a prognostic indicator.

### Background of the Invention

25 In recent years there has been significant progress in deciphering the molecular changes accompanying malignant transformation. A large number of tumors have been characterized as carrying a variety of chromosomal and submicroscopic genomic modifications, including activation of dominant-acting proto-oncogenes, inactivation of tumor suppressor genes and inactivation of  
30 metastasis suppressor genes. It is presumed that such tumors are descended from cell lineages that have accumulated a set of critical genetic lesions through random and relatively rare mutations at defined chromosomal locations. As an alternative tumor mechanism, some have postulated that a generalized,

disseminated genomic instability, with attendant increased frequency of mutations at numerous unrelated loci, may represent a primary genetic mechanism in some tumors.

In a normal cell, multiple metabolic pathways control the overall accuracy of various functions, including DNA replication and repair, cell division and differentiation. A generalized genomic instability potentially could arise from disruption of one or more of these functions, including DNA replication, post-replicative proofreading, DNA repair, cell cycle checkpoint proteins, and DNA recombination. Mutations in genes that cause a generalized increase in the frequency of substitutions, insertions, deletions or other structural changes throughout the genome can be classified as "mutator" mutations.

Several clinical correlates have been cited in support of the theory that mutations in "mutator genes" may alter the regulation of a wide spectrum of genes, including those genes responsible for tumorigenesis. For example, individuals with the inherited disorder xeroderma pigmentosum are defective in excision repair of DNA. A clinical correlate of this condition is a predisposition of these individuals to skin cancer following exposure to ultraviolet light.

In summary, malignant transformation may involve either of the following genetic pathways:

(1) accumulation in selected cell lineages of random but relatively infrequent mutations in proto-oncogenes, tumor suppressor genes and other genes directly related to tumorigenesis.

(2) mutation(s) in a mutator gene with resultant genomic instability.

There may be functional overlap in these two mechanisms in the sense that mechanism (2) may lead secondarily to mutations in known proto-oncogenes and tumor suppressor genes. However, the generalized and disseminated genomic instability of mechanism (2) may influence carcinogenesis through a wide array of known and unknown genetic mechanisms. As such, there may be little or no correlation between elevated levels of genomic

instability and known mutational changes in characterized tumor-related genes. Moreover, it is conceivable that patients having tumors deriving from mechanism (2) may enjoy a relatively favorable prognosis compared to patients having tumors deriving from mechanism (1). This is due to the tendency of tumor cells having a generalized and disseminated genomic instability to continue to accumulate mutational changes, leading to cell disfunction and/or death. That is, these tumors may be relatively self-limiting in comparison to tumors deriving from mechanism (1).

While a variety of nucleic acid probe-based assays are available to detect structural alterations in known tumor-related genes, there has been no reliable indicator of mechanisms involving generalized genomic instability in tumor formation. Nor has there been a reliable test to distinguish tumors arising from random mutations in tumor-related genes from those tumors arising from a more generalized genomic instability. The present inventors have discovered that monitoring the structural integrity, or sequence organization, of one or more members of selected interspersed repetitive sequence families provides a method for assessing genomic instability in tumors and for correlating such instability with tumor location and clinical prognosis.

One form of interspersed repetitive sequence in humans is the "microsatellite" class of sequences. Microsatellites are generally composed of about one to six base-pair motifs that can be tandemly repeated multiple times. Microsatellite sequences of the form  $(CA)_n \cdot (GT)_n$  constitute one of the most abundant classes of repetitive DNA families in human DNA. Approximately 50,000-100,000  $(CA)_n$ -repeats are scattered throughout the human genome. Many, if not most, of these  $(CA)_n$ -repeats exhibit length polymorphisms. Dracopoli et al., Mapping the Human Amylase Gene Cluster on the Proximal Short Arm of Chromosome 1 Using a Highly Informative  $(CA)_n$  Repeat, Genomics 7:97-102 (1990); Nurnberg et al., DNA Fingerprinting With the Oligonucleotide Probe  $(CAC)_5/(GTG)_5$ : Somatic Stability and Germline Mutations, Hum. Genet. 84:75-78 (1989); Tautz, Hypervariability of Simple Sequences as a General Source for Polymorphic DNA Markers, Nucleic Acids Res. 17:6463-71 (1989). Although the rate of new mutations at these sites is slightly increased compared to other

genomic sites, the overall rate is still quite low. In general, alleles at these sites are stably inherited from one generation to another. In fact, these markers have now been identified as one of the most useful classes of DNA polymorphism for the purpose of linkage analysis. Weissenbach et al., A Second-Generation  
5 Linkage Map of the Human Genome, Nature 359: 794-801 (1992).

Although specific forms of genomic instability previously have not been correlated with specific tumors or with tumor locations and tumor prognoses, probes directed toward "minisatellite" interspersed repetitive sequence families have detected novel fragments in various malignancies. In  
10 comparison to microsatellites, minisatellites are characterized by more complex short repetitive stretches of noncoding sequences. Moreover, minisatellite loci tend to be clustered on the tips of chromosome arms. Weissenbach et al., at 794.

Southern blot analysis of genomic DNA with probes consisting of minisatellite sequences reveals an array of fragments that vary in molecular  
15 weight. Each fragment represents multiple non-contiguous loci within the human genome. As such, these probes have been useful for multilocus fingerprinting of DNA. Jeffreys et al., Individual-Specific Fingerprints of Human DNA, Nature 316: 76 (1985); Jeffreys et al., Minisatellite Repeat Coding as a Digital Approach to DNA Typing, Nature 354: 204 (1991). In a similar fashion,  
20 alterations in multilocus microsatellite sequences of the form  $(CA)_n$  may be detected in genomic Southern blots using probes comprising various  $(CA)_n$  nucleotide stretches. However, a more rigorous investigation of mutational changes in individual microsatellite loci may be obtained by amplification of individual microsatellites using primer pairs directed to unique sequence  
25 flanking such microsatellites.

Additional bands or deleted bands have been observed in genomic blots of tumor DNA probed with  $(CAC)_5$  and  $(GTG)_5$  oligonucleotide probes. Lagoda et al., Increased Detectability of Somatic Changes in DNA from Human Tumours After Probing with "Synthetic" and "Genome-Derived" Hypervariable  
30 Multilocus Probes, Hum. Genet. 84: 35-40 (1989). However, such changes were not correlated with genomic instability as related to particular tumor types, with particular anatomical groupings of tumors, or with clinical prognosis.

The sequences detected by minisatellite and microsatellite probes are not known to have any specific function. Thus, the abnormalities detected are not expected, in themselves, to be causally involved in tumor initiation or progression. However, heritable, unstable DNA elements recently have been identified as the basis of disease for three separate inherited disorders: (1) Fragile X syndrome, Kremer et al., Mapping of DNA Instability at the Fragile X to a Trinucleotide Repeat Sequence p(CCG)<sub>n</sub>, Science 252:1711-14 (1991); Fu et al., Variation of the CGG Repeat at the Fragile X Site Results in Genetic Instability: Resolution of the Sherman Paradox, Cell 67:1047-58 (1991); Hirst et al., Genotype Prediction in the Fragile X Syndrome. J. Med. Genet. 28:824-29 (1991); (2) Kennedy's disease, La Spada et al., Androgen Receptor Gene Mutations in X-Linked Spinal and Bulbar Muscular Atrophy, Nature 352:77-79 (1991); and (3) Myotonic Dystrophy, Mahadevan et al., Myotonic Dystrophy Mutation: An Unstable CTG Repeat in the 3' Untranslated Region of the Gene, Science 255:1253-55 (1992); Brook et al., Molecular Basis of Myotonic Dystrophy: Extension of a Trinucleotide (CTG) Repeat at the 3' End of a Transcript Encoding a Protein Kinase Family Member, Cell 68:799-808 (1992). All available evidence suggests that amplification of a tri-nucleotide repeat is involved in the molecular pathology in each of these disorders. Although these tri-nucleotide repeats appear to be in non-coding DNA, they clearly are involved with perturbations of genomic regions that ultimately affect gene expression. Perturbations of various di- and tri-nucleotide repeats resulting from somatic mutation in tumor cells could also affect gene expression and/or gene regulation.

To investigate the role of disseminated genomic instability as a mechanism in tumorigenesis, it is useful to examine model systems in which tumors of potentially distinctive genetic backgrounds may be identified. Preliminary studies have indicated that different genetic mechanisms of tumorigenesis may be operative in different anatomical regions of the colorectal tract. On the one hand, evidence indicates that the process of tumorigenesis in at least some colorectal cancers proceeds through a series of genetic alterations in defined loci including both dominant and recessive acting proto-oncogenes. On the other hand, some colorectal cancers do not display such defined structural

changes in known tumor-associated genes. Colorectal cancer therefore represents a useful model for analysis of disseminated genomic instability as a potential mechanism in cancer.

With respect to dominant acting proto-oncogenes in colorectal cancer, both carcinomas and the larger villous type of adenomas have shown point mutation in the ras proto-oncogene in roughly 50% of the cases. Vogelstein et al., Genetic Alterations During Colorectal-Tumor Development, N. Engl. J. Med. 319:525-32 (1988). C-myc mRNA has been found to be expressed at significantly higher levels in tumors compared to normal mucosa. Sikora et al., C-myc Oncogene Expression in Colorectal Cancer, Cancer 59:1289-95 (1987); Erisman et al., The C-myc Protein is Constitutively Expressed at Elevated Levels in Colorectal Carcinoma Cell Lines, Oncogene 2:367-78 (1988).

With respect to recessive acting proto-oncogenes (e.g., tumor-suppressor genes), several studies have demonstrated allelic loss in colorectal carcinoma. Loss of allelic heterozygosity (LOH) on a particular chromosome in cells of a tumor provides indirect evidence for the presence of a tumor suppressor gene(s) on that portion of the chromosome involved in the LOH. It has been proposed that tumor suppressor gene function can be lost through mutational inactivation of one member of an allelic pair and an accompanying chromosomal deletion that leads to physical loss or inactivation of the other member of the allelic pair. The chromosomal deletion is detected as an LOH using a polymorphic marker present in the same chromosomal region as the tumor suppressor gene. The most frequent sites of allelic loss in colon cancer appear to be on chromosomes 17 and 18 (each exhibiting LOH in nearly 75% of colorectal carcinomas); another one-third to one-half of such tumors exhibit LOH on chromosome 5. Candidate tumor suppressor genes are MCC and APC (familial adenomatous polyposis locus) on chromosome 5q, p53 on chromosome 17p and DCC (deleted in colorectal carcinoma) on chromosome 18q. Kinzler et al., Identification of a Gene Located at Chromosome 5q21 That Is Mutated in Colorectal Cancer, Science 251:1366-70 (1991); Kinzler et al., Identification of FAP Locus Genes from Chromosome 5q21, Science 253:661-64 (1991); Baker et al., Chromosome 17 Deletions and p53 Gene Mutations in Colorectal Carcinomas,



Science 244:217-21 (1989), Fearon et al., Identification of a Chromosome 18q Gene That Is Altered in Colorectal Cancers, Science 247:49-56 (1990). In addition to chromosomes 5, 17 and 18, other chromosomes including chromosomes 1, 6, 8, 14 and 22 have been implicated in the genesis of colorectal cancer.

5           To date, results of LOH studies in colorectal cancer have shown a significant correlation with the site of the tumor. Delattre et al., Multiple Genetic Alterations in Distal and Proximal Colorectal Cancer, The Lancet 2:353-56 (1989); Kern et al., Allelic Loss in Colorectal Carcinoma, J. Am. Med. Assoc. 261:3099-3103 (1989); Offerhaus et al., The Relationship of DNA Aneuploidy to Molecular  
10 Genetic Alterations in Colorectal Carcinoma, Gastroenterology 102:1612-19 (1992). Specifically, it appears that allelic loss on chromosomes 5, 17 and 18 occurs more frequently in distal tumors than in proximal tumors. Such differences suggest that proximal and distal tumors may arise through different pathogenetic mechanisms, and that at least two genetic categories of colorectal  
15 cancer may exist. In fact, a growing body of evidence suggests that tumors of the proximal colon differ from tumors of the distal colon. Bufill et al., Colorectal Cancer: Evidence for Distinct Genetic Categories Based on Proximal or Distal Tumor Location, Ann. Int. Med. 113:779-88 (1990). This evidence, which stems from developmental and biological differences within the normal colon and  
20 various characteristics noted in colorectal cancer, includes the following:

A.   Developmental. The area extending from the cecum to the proximal two-thirds of the transverse colon is derived from the embryonic mid gut while the distal third of the transverse colon to the rectum is derived from the embryonic hind gut. Langman, J.,  
25 Medical Embryology, 4th Ed. (1981). The distinct embryologic origin of these two regions is also reflected in different vascular supplies. Subsequent development leads to a number of properties that differ between the proximal and distal colon, such as the distribution of hormone producing cells.

30       B.   Biological. Biological differences between proximal and distal colon include differences in ability to metabolize various carcinogens, differences in the expression of cell surface antigens,

and differences in enzymatic activities. Stralka et al., Cytochrome P-450 Activity and Distribution in the Human Colon Mucosa, Cancer 64:2111-16 (1980); Hughes et al., Antigen Expression in Normal and Neoplastic Colonic Mucosa: 3 Tissue-Specific Antigens Using Monoclonal Antibodies to Isolated Colonic Glands, Cancer Res. 46:2164-71 (1986); Tari et al., The Relation Between Ornithine Decarboxylase Activity and the Location of Colorectal Cancers, Gastroenterology 98:A313 (1990).

C. Epidemiological. The relative incidence of proximal and distal tumors appears to vary in relation to the total incidence in a given region. In regions with a low incidence, proximal tumors predominate. In high risk areas, distal tumors predominate. Furthermore, the incidence of proximal tumors in the population at large appears to be increasing. Ghahremani et al., Colorectal Carcinomas: Diagnostics Implications of their Changing Frequency and Anatomic Distribution, World J. Surg. 13:321-25 (1989); Fleshner et al., Age and Sex Distribution of Patients with Colorectal Cancer, Dis. Rectum 32:107-11 (1989).

D. Inheritance. Colon cancer that appears in patients with familial adenomatous polyposis tends to be left-sided while those that occur in hereditary non-polyposis rectal cancer tend to be right-sided. Lynch et al., Hereditary Non-Polyposis Rectal Cancer (Lynch Syndromes I and II). II. Biomarker Studies, Cancer 56:939-51 (1985).

E. Other Proto-oncogenes. Several studies have shown that tumors expressing deregulated c-myc are found more frequently in the distal colon. Rothberg et al., Evidence that C-myc Expression Defines Two Genetically Distinct Forms of Colorectal Adenocarcinoma, Br. J. Cancer 52:629-32 (1985); Sikora et al., C-myc Oncogene Expression in Colorectal Cancer, Cancer 59:1289-95 (1987); Erisman et al., The C-myc Protein is Constitutively Expressed at Elevated Levels in Colorectal Carcinoma Cell Lines, Oncogene 2:367-78 (1988). On the other hand, mutations in the ras proto-

oncogene occur more frequently in proximal tumors. Offerhaus et al., The Relationship of DNA Aneuploidy to Molecular Genetic Alternations in Colorectal Carcinoma, Gastroenterology 102:1612-19 (1992).

5           F. Tumor DNA Content. Analysis of DNA content by flow cytometry has shown that aneuploidy is more frequently encountered in distal tumors than in proximal tumors. Delattre et al., Multiple Genetic Alternations in Distal and Proximal Colorectal Cancer, Lancet 2:353-56 (1989).

10           All of the above-referenced differences suggest the tendency of distinct pathogenetic mechanisms to be responsible for cancers of the proximal and distal colon. Although tumor suppressor genes appear to play a more prominent role in the development of tumors in the distal as opposed to the proximal colon, the underlying genetic lesions in proximal tumors have been  
15           poorly understood. Nevertheless, knowledge of the molecular genetic mechanisms involved in cancer will have an enormous impact on clinical practice and on the design of future prospective research studies covering new methods of surgical adjuvant therapy. There is an ongoing need for new prognostic indicators to more precisely identify those patients that may benefit  
20           from adjuvant therapy and those patients at high risk for tumor recurrence.

          Another potential model for examination of disseminated genomic instability is the so-called Muir-Torre syndrome (MTS). MTS is defined by the development in an individual of at least one sebaceous gland tumor and a minimum of one internal malignancy. Muir, E.G., Bell, A.J.Y., and Barlow, K.A.,  
25           Multiple Primary Carcinomata of the Colon, Duodenum, and Larynx Associated with Kerato-acanthoma of the Face, Br. J. Surg. 54:191-195, 1967; Torre, D., Multiple Sebaceous Tumors, Arch. Dermatol. 98:549-551, 1968. Although a wide range of internal malignancies has been reported, the most frequently observed neoplasm is colorectal carcinoma. Finan, M.C., and Connolly, S.M., Sebaceous  
30           Gland Tumors and Systemic Disease: A Clinicopathologic Analysis, Medicine 63:232-242, 1984; Cohen, P.R. et al., Association of Sebaceous Gland Tumors and Internal Malignancy: The Muir-Torre Syndrome, Am. J. Med. 90:606-613, 1991.

Interestingly, MTS shares a number of clinical and pathologic features with hereditary nonpolyposis colorectal cancer (HNPCC) syndromes Lynch I and Lynch II, including: 1) an autosomal dominantly inherited susceptibility to colorectal cancer; 2) onset of colorectal cancer at any early stage; 3) tendency to develop right sided colon cancers; and 4) prolonged survival after diagnosis of colorectal cancer. Lynch, H.T. et al., Genetics, Natural History, Tumor Spectrum, and Pathology of Heredity Nonpolyposis Colorectal Cancer: An Updated Review, Gastroenterology 104:1535-1549, 1993. In addition, MTS shares with Lynch II a predisposition for the development of extra-colonic tumors, especially those of the genitourinary tract. MTS differs from Lynch II syndrome, however, in the development of sebaceous tumors of the skin. As a result of the similarities noted above, Lynch has suggested that the Muir-Torre syndrome may be a fuller phenotypic expression of the gene responsible for HNPCC. Lynch, H.T. et al., The Cancer Family Syndrome: Rare Cutaneous Phenotypic Linkage of Torre's Syndrome, Arch. Intern. Med. 141:607-611, 1981; Lynch, H.T. et al., Muir-Torre Syndrome in Several Members of a Family with a Variant of the Cancer Family Syndrome, Br. J. Dermatol. 113:295-301, 1985.

#### Summary of the Invention

The present inventors have discovered a method for detecting genomic instability in malignant tumors through analysis of interspersed repetitive sequences in the DNA of such tumors. The analysis can provide a prognostic indicator for patients having such tumors. The method involves obtaining genomic DNA from a tumor and from normal tissue of the same patient and evaluating one or more interspersed repetitive sequence loci in the respective DNAs. The structural status, or sequence organization, of the interspersed repetitive sequences in the tumor is compared with the structural status of the interspersed repetitive sequences in the corresponding normal tissue. The results of this analysis may be correlated with the presence or absence of genomic instability in the tumor genome and with the clinical prognosis of the patient having the tumor.

The interspersed repetitive sequences examined may be minisatellite or microsatellite sequences. For example, various di- or tri-nucleotide microsatellite repeats may be examined at one or more loci. In particular, a microsatellite di-nucleotide repeat of the form  $(CA)_n \bullet (GT)_n$  has been  
5 found useful for analysis of colorectal tumors. In one embodiment, selected di- and tri-nucleotide repeat loci are examined using sequence amplification procedures. Appropriate primers for polymerase-based amplification of selected loci are chosen and used to amplify, for example, one or more selected  $(CA)_n$  repeat loci. Polymerase chain reaction or other amplification methods may be  
10 used in the analysis. Primer pairs may be selected to hybridize to unique sequence flanking the 3' and 5' ends of the  $(CA)_n$  repeat region of the locus. The apparent fragment sizes of the amplified DNA's are compared in order to evaluate whether or not the  $(CA)_n$  repeat region has undergone a structural alteration (e.g., insertion/amplification, deletion) in the tumor DNA that is not  
15 seen in the corresponding normal DNA.

The presence of structural alterations in one or more interspersed repetitive sequence loci in a tumor is interpreted as evidence for a generalized, disseminated genomic instability. The structural alterations can be correlated with the clinical prognosis of the patient having the tumor.

20

#### Brief Description of the Figures

Figure 1 depicts autoradiographs of amplified fragments obtained with primer pairs specific for  $(CA)_n$  di-nucleotide repeat microsatellite loci.

Figure 2 depicts an autoradiograph of amplified fragments obtained  
25 with primer pairs specific for a tri-nucleotide repeat microsatellite locus.

Figure 3 depicts an autoradiograph of a sequencing gel containing the products of sequencing reactions carried out on amplified fragments obtained with the Mfd 41 primer pair.

Figure 4 depicts an autoradiograph of a sequencing gel containing  
30 the products of sequencing reactions carried out on amplified fragments obtained with the Mfd 27 primer pair.

Figure 5 depicts an analysis of DNA from paired normal (N) and tumor (indicated by arabic number) tissue for the presence of microsatellite instability at two loci, P53 (panel A) and 635/636 (panel B). Case numbers are shown above the lanes. The + and - signs indicate whether tumors were scored positive or negative for microsatellite instability. Tumors shown include: a) a cecal adenocarcinoma (1) and a keratoacanthoma (2) from case 4; b) a sebaceous carcinoma (1), sebaceous adenoma (2), colonic adenocarcinoma (3), and a lymph node diagnosed with CLL (4) from case 9; and c) a Meibomian gland adenoma (1) and breast carcinoma (2) from case 19.

#### Detailed Description of the Invention

The present inventors have discovered a type of genetic alteration appearing in malignant tumor cell DNA and not in normal cell DNA. These alterations are interpreted to be indicative of a disseminated, generalized form of genomic instability associated with at least one mechanism of tumorigenesis. This form of genomic instability may be detected through examination of interspersed repetitive DNA sequences. Preferably these sequences are members of the "minisatellite" and "microsatellite" interspersed repetitive sequence families. Most preferably these sequences are members of the microsatellite interspersed repetitive sequence families.

In order to characterize alterations at these sites in detail, DNA from colorectal cancers or other tumors may be examined using gene amplification-based assays for individual loci comprising the (CA)<sub>n</sub> form of microsatellite interspersed repetitive sequence. Loci selected for such analysis may be selected for proximity to known dominant acting proto-oncogenes, tumor suppressor genes or chromosome locations otherwise suspected of being involved in genetic changes accompanying malignant transformation. Alternatively, other loci may be chosen at relatively random locations on various chromosomes. Tumor-specific changes in the apparent fragment sizes of such alleles can vary from single 2 base-pair (bp) alterations to larger insertions/amplifications or deletion events. For purposes of data analysis of genomic instability, the present inventors have classified the larger insertion/amplification/deletion events as

"Type I" mutations and the smaller (e.g. 2 bp) alterations as "Type II" mutations. Other classifications of DNA structural changes in various tumor systems may be appropriate depending on the array of mutational changes in interspersed repetitive sequences observed in such tumors.

5               For examination of generalized genomic instability in tumors, tumor samples may be removed from patients diagnosed with malignant tumors. Tumor tissue may be obtained through biopsies or following surgical removal of all or part of the tumor. DNA from normal tissue of such patients may be obtained from peripheral white blood cells, from normal tissue infiltrates  
10 of the tumors, from biopsy or surgical removal of normal tissue adjacent to the tumors or by other methods known to those skilled in the art. DNA may be obtained from unfixed or frozen tissue or, alternatively, from tissue prepared for histology. For example, DNA may be obtained from tumor tissue embedded in paraffin.

15               In a preferred embodiment, tumor and tissue samples are processed to ensure that the tumor samples contain an operable fraction of tumor cells and that the normal tissue samples are operably free of tumor cells. In the experiments reported in the Examples below, the tumor samples were processed to have at least 70% tumor cells. The primary consideration in processing of  
20 tumor tissue is that a sufficient, or operable, fraction of tumor tissue is present to allow detection of DNA structural changes in tumor DNA samples compared to normal tissue DNA samples. Although 70% tumor cells was chosen as the minimum fraction of tumor cells in the experiments reported in the Examples, below, other percentages may be appropriate as long as nucleic acid sequence  
25 changes are correctly interpretable. The normal tissue sample should be operably free of tumor cells, such that any contaminating tumor cells do not cause a misreading of the normal band positions. Preferably the normal tissue sample is absolutely free of tumor cells. Operable fractions of cells for a given tumor system may be readily determined through routine preliminary experiments.  
30 For example, relatively pure samples of normal and tumor cells may be mixed in various proportions and the resulting mixtures used to provide DNA samples.

Results from the mixed samples are then readily evaluated for accuracy of results compared to the pure samples.

DNA may be isolated from fresh or frozen tissue using standard methods as described in, for example, Sambrook et al., Molecular Cloning: A Laboratory Manual, 2d Ed. (1989), pages 9.16-9.23. Isolation of DNA from paraffin-embedded tissues is described, for example, by Jackson et al., Tissue Extraction of DNA and RNA and an Analysis by the Polymerase Chain Reaction, J. Clin. Pathol. 43:499-504 (1990) and Rogers et al., Analysis of DNA in Fresh and Fixed Tissue by the Polymerase Chain Reaction, Am. J. Pathol. 136:541-48 (1990).

Once the genomic DNA is isolated from the tumors and normal tissues, polymerase chain reactions or other gene amplification reactions may be carried out using standard methods. Analysis of microsatellite sequences by polymerase chain reaction, for example, may follow the methods of Weber and May, Abundant Class of Human DNA Polymorphisms Which Can Be Typed Using the Polymerase Chain Reaction, Am. J. Hum. Genet. 44:388-96 (1989). Primer pairs are selected to hybridize to the DNA flanking the interspersed repetitive sequence loci at selected chromosomal locations. Microsatellite alleles may be chosen, for example, for their proximity to regions of chromosomes that have been implicated in allelic loss in various tumors. Conversely, other alleles may be selected specifically for chromosomal locations known to have a low frequency of allelic loss in colorectal or other cancers.

Aliquots of the amplified DNA's may be mixed with formamide sample buffer and electrophoresed on standard denaturing polyacrylamide DNA sequencing gels. Aliquots of the tumor and tissue DNA from the same patient and amplified using the same primer pair are loaded next to one another in the gel. Gels are then fixed, dried and processed for autoradiography. The lanes containing amplified tumor and tissue DNA are compared to one another to determine the mutational status of the amplified tumor DNA at the locus tested.

A variety of gene amplification techniques may be used for analysis of individual loci of interspersed repetitive sequences. Such methods may include, without limitation, Polymerase Chain Reaction (PCR), Saiki et al., Enzymatic Amplification of Beta-Globin Genomic Sequences and Restriction Site



Analysis for the Diagnosis of Sickle Cell Anemia, Science 230:1350-54 (1985); Ligase Chain Reaction (LCR), Wu and Wallace, The Ligation Amplification Reaction (LAR) - Amplification of Specific DNA Sequences Using Sequential Rounds of Template-Dependent Ligation, Genomics 4:560-69 (1989) and  
5 Landegren et al., A Ligase Mediated Gene Detection Technique, Science 241:1077-80 (1988); Q-beta-Replicase Template Amplification (Q-beta), Lomeli et al., Quantitative Assays Based on the Use of Replicable Hybridization Probes, Clin. Chem. 35:1826-31 (1989); and Strand Displacement Activation (SDA), Walder et al., Isothermal *in vitro* Amplification of DNA by a Restriction Enzyme/DNA  
10 Polymerase System, Proc. Nat'l Acad. Sci. USA 89:392-96 (1992). RNA-based amplification methods may be used for those interspersed repetitive sequences that are expressed as RNA. An example of an RNA-based amplification method is Self-Sustained Sequence Replication (3SR), Guatelli et al., Isothermal In Vitro Amplification of Nucleic Acids by a Multi-enzyme Reaction Modified After  
15 Retroviral Replication, Proc. Nat'l Acad. Sci. USA 87:1874 (1990).

As an alternative to analysis of interspersed repetitive sequence loci using sequence amplification methods, genomic instability also may be examined through analysis of the total population of selected microsatellite and/or minisatellite DNA families. For example, sequences comprising (CA)<sub>n</sub>  
20 oligonucleotides may be used as probes of genomic southern blots to assess the overall structural integrity of the (CA)<sub>n</sub> interspersed repetitive sequence family. Such "whole-family" scanning, although less discriminating of individual loci than the amplification-based analysis described above, may be useful for particular tumor systems.

25 The present inventors have discovered that structural alterations of interspersed repetitive sequences, as sign posts of generalized genomic instability, may be correlated with identifiable tumor sub-groupings and clinical prognosis. For example, differences in genomic instability as revealed by alterations in interspersed repetitive sequences have been found by the present inventors to  
30 vary with anatomical location of tumors within the colorectal region and with disease recurrence and survival. As discussed above, such sublocations or

subgroupings of primary tumors may tend to arise by distinct genetic mechanisms.

Similar phenomena also may be expected to occur in other organ systems. For example, although not providing correlations with anatomical subgroupings of primary tumors or with clinical prognosis, several studies have demonstrated alterations in interspersed repetitive sequences in diverse tumor systems including stomach carcinoma as discussed by Lagoda et al., Increased Detectability of Somatic Changes in DNA from Human Tumours After Probing with "Synthetic" and "Genome-Derived" Hypervariable Multilocus Probes, Hum. Genet. 84: 35-40 (1989); breast, adrenal, pancreas and thyroid tumors as discussed by Matsumura, Y. and D. Tarin, DNA Fingerprinting Survey of Various Human Tumors and Their Metastases, Cancer Res. 52: 2174-2179 (1992). Thus, the present inventors' discovery of an association of specific tumor characteristics with variations in genomic stability of colorectal tumors may be replicated with a similar analysis of interspersed repetitive sequences in other organ tumor systems.

In this regard, the present inventors have examined microsatellite instability (MIN) in tumors of patients diagnosed with MTS. Results of the present experiments suggest that the Muir-Torre syndrome includes at least two subgroups of patients: one subgroup exhibiting the clinical features of Lynch II syndrome and demonstrating MIN; and the other not exhibiting the features of Lynch II syndrome and not demonstrating MIN. The present inventors have detected MIN in diverse and seemingly unrelated tumors from MTS patients (e.g. colorectal carcinoma, transitional cell carcinoma and sebaceous adenoma), and have also identified both MIN+ and MIN- tumors in the same MTS patient. It is possible that the tissues from which these MIN+ neoplasias arise are more susceptible than tissues of MIN- tumor origin to a "mutator gene effect" because of the influence of environmental factors, such as UV exposure or various carcinogens. Alternatively, other tissue specific modifier genes may play a role in neoplasia of these tissues. In any case, it appears that each of the tumors demonstrating MIN in MTS shares a common underlying molecular basis of carcinogenesis.

The indolent nature of the tumors in MTS and HNPCC is one of the most intriguing features of these syndromes. Prolonged survival of patients whose tumors were MIN+ was observed despite the fact that the majority of these patients developed metachronous colon cancers, some of advanced stage.

5 For instance, case 24 was a female who developed a cecal carcinoma (modified Dukes B1) at age 23, a metachronous rectal adenocarcinoma (modified Dukes B2) at age 29, metastatic grade 1 mucinous adenocarcinoma (modified Dukes D) to her liver at age 30, recurrent colon carcinoma at a previous anastomotic site at age 51 and a second rectal cancer at age 64. Each of these tumors was resected and  
10 presently at age 70 she is alive and without evidence of cancer. Similarly, cases 6, 9, 23 and 25 experienced long durations of survival despite the development of metachronous colon cancers and other visceral malignancies.

It is unclear why patients having tumors with MIN exhibit prolonged survival. Tumors with widespread MIN may contain mutations in a  
15 gene or genes that, when defective, promote instability at numerous loci, including those defined by microsatellites. Such mutations could non-specifically alter the regulation of a wide spectrum of genes, thereby promoting tumor formation. By the same mechanism, such mutations could also disrupt the normal function of genes that are critical to cell vitality and viability and  
20 thereby reduce the biological aggressiveness of the tumor. Alternatively, the intrinsic phenotype of malignancies associated with MIN may have an attenuated tumorigenic or metastatic potential. Further clarification, however, must await characterization of the gene(s) responsible for the particular phenotype.

25 Despite the apparently good prognosis associated with widespread MIN in MTS, it is important to note that two patients in the MIN+ group of MTS patients developed malignancies (a renal cell carcinoma in case 6 and a chronic lymphocytic leukemia in case 9) that did not demonstrate MIN. Unlike tumors demonstrating microsatellite instability, these cancers did not behave indolently  
30 and eventually led to the patients' deaths. The absence of MIN in these malignancies argues that some of the tumors that occur in MTS patients could arise through separate genetic mechanisms. Thus, while patients with MTS are

5           The invention will be further understood with reference to the following illustrative embodiments, which are purely exemplary, and should not be taken as limiting the true scope of the present invention as described in the claims.

### EXAMPLE 1

## Isolation of Tumor and Normal Tissue DNA and Analysis for Genomic Instability

DNA from 90 colorectal cancers and the corresponding DNA from normal tissue was isolated from 87 patients in the following manner. A hematoxylin and eosin-stained frozen section of each tumor was examined and used as a template to dissect away as much non-tumor tissue as possible in order to enrich for tumor cells. The remaining tumor fragments were remounted for serial frozen sections in Tissue-Tek ornithine carbamyl transferase compound (Miles, Inc., Cat. No. 4583). One hundred 10-20 micron-thick cryostat sections were then made. The first, fiftieth, and one-hundredth sections were stained with hematoxylin and eosin, and the fraction of tumor nuclei present at these levels was assessed by a pathologist. Only those tumors containing  $\geq 70\%$  tumor nuclei in each of these three sections were used for subsequent steps. DNA was extracted from tumor and normal control tissue (either normal mucosa or peripheral blood leukocyte) using an Applied Biosystem 340A Nucleic Acid Extractor (ABI, Inc.) in accordance with the manufacturer's instructions and stored at  $4^{\circ}\text{C}$  until used.

Once the genomic DNA was isolated from the tumors and tissues and prepared as described above, standard polymerase chain reactions were carried out following the methods of J.L. Weber and P.E. May, Abundant Class of Human DNA Polymorphisms Which Can Be Typed Using the Polymerase Chain Reaction, *Am. J. Hum. Genet.* 44:388-396 (1989). The isolated genomic

DNA was tested with four primer pairs that hybridize to the flanking DNA of interspersed repetitive sequence loci in the genome. The loci included (CA)<sub>n</sub>-repeats on chromosome 5 (Mfd 27), chromosome 17 (Mfd 41), chromosome 18 (Mfd 26), and chromosome 15 (635). The primer pairs are set forth in Table 1 and

5 in the Sequence Listing as SEQ. ID. NOS. 1-8.

All polymerase chain reactions were performed in a Perkin-Elmer Cetus Thermal Cycler. Reaction were carried out in a 25 µl reaction volume comprising:

	5.0	µl	5X Reaction Buffer Containing dNTPs
10	1.0	µl	First Primer (20 µM)
	1.0	µl	Second Primer (20 µM)
	1.0	µl	Genomic DNA (250 µg/ml)
	0.25	µl	Taq-I polymerase (Perkin Elmer Cetus, 5 units/µl)
	0.25	µl	<sup>32</sup> P alpha-dCTP (3000 Ci/mmol)
15	16.5	µl	sterile H <sub>2</sub> O

The formula for 5X Reaction Buffer Containing dNTPs is as follows:

	1 mM dATP
20	1 mM dCTP
	1 mM dGTP
	1 mM dTTP
	250 mM KCl
	50 mM Tris-Cl, pH 8.3
25	7.5 mM MgCl <sub>2</sub>
	0.05% (w/v) gelatin

Samples were kept on ice during mixing and then processed through appropriate temperature cycles for amplification. A typical

30 amplification procedure comprised 30 temperature cycles each consisting of 1 minute at 94°C (denaturation) and 2 minutes at 55°C (annealing). Generally, the elongation step of the last cycle was lengthened an additional 10 minutes. Routine variations of these parameters were employed as necessary to optimize the amplification reactions with individual primer pairs. Following this

35 temperature cycling, the reaction volume was brought to 250 µl with H<sub>2</sub>O. Aliquots of the diluted, amplified DNA were mixed with 2 volumes loading buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF), heated at 90° C for 2 minutes to denature the DNA, and

electrophoresed on standard denaturing polyacrylamide DNA sequencing gels. Aliquots of the tumor and tissue DNA from the same patient and amplified using the same primer pair were loaded next to one another in the gel. Gels were then dried and processed for autoradiography. Autoradiograph exposure  
5 times were about 2 days at room temperature.

After the autoradiographic film was developed, the lanes containing amplified tumor and tissue DNA were compared to one another to determine the mutational status of the amplified tumor DNA at the locus tested. The film was examined in order to identify whether the banding pattern varied  
10 between the tumor and tissue lanes. Generally, multiple bands were detected for each allele: one major band and multiple shadow bands. The shadow bands are thought to be due to PCR artifact (i.e., strand slippage). As shown in Figure 1, significant variability of fragment sizes reflecting genomic instability was detected in tumor specimens, ranging from a 2 base-pair change in some (Panel  
15 B, lane 15 compared to lane 14) to a larger insertion-amplification (panel B, lane 2 compared to lane 1) or deletion in others (Panel A, lane 9 compared to lane 8). Alterations intermediate to these are shown in Panel B, lanes 3 and 18. For purposes of data analysis, two general types of genomic instability were defined. A significant increase (insertion/amplification) or decrease (deletion) in the  
20 apparent fragment size compared to the alleles found in DNA from the normal tissue were called Type I mutations (Panel A, lanes 4, 5, 7 and 9; Panel B, lanes 2, 3, 5, 6, 18 and 19). Single two base-pair alterations in the apparent fragment size were called Type II mutations (Panel B, lanes 15 and 16).

Genetic alterations in these microsatellite interspersed repetitive  
25 sequence loci were detected in 25 (28%) of the 90 tumors examined. Eight of these 25 positive tumors contained only Type I mutations, 12 contained only Type II mutations, and 5 showed the presence of both Type I and Type II mutations. All 13 tumors that possessed a Type I mutation demonstrated alterations at multiple loci. Of these 13 tumors, nine showed alterations within  
30 all (CA)<sub>n</sub>-repeats examined (Table 2). Tumors having Type II mutations typically showed a change at only a single locus (10 of 17 tumors). Of the 12 tumors with

only Type II mutations, ten showed an alteration at only a single (CA)<sub>n</sub>-repeat (Table 2).

No such genetic alterations were found in any of the paired normal tissues examined.

5

## EXAMPLE 2

### Correlations With Loss of Heterozygosity (LOH)

Fourteen probes (see Table 3) detecting DNA polymorphisms on chromosomes 5, 17 and 18 were used to assess the frequency of allelic loss in a group of 91 colorectal cancers. Allelic loss was defined as the demonstration of loss of heterozygosity by one or more of the probes on that chromosome or chromosomal arm.

Tumor and normal tissues were processed and DNA isolated according to the procedures set out in Example 1. The DNA from the tumors and tissues was digested with an appropriate restriction endonuclease, electrophoresed in an 0.8% agarose gel and transferred to a nylon filter from the agarose gel using the method of Southern. Sambrook et al., Molecular Cloning: A Laboratory Manual, 2d Ed. (1989), pages 9.31-9.62. The DNA was subsequently hybridized with chromosome-specific <sup>32</sup>P-labelled DNA probe. The gene probes used and their chromosomal positions are shown in Table 3. Hybridization with the labelled probe was detected by autoradiography.

For the 91 tumors analyzed, chromosomes 17 and 18 showed the highest frequency of allelic loss (69% and 70% of tumors, respectively), while a frequency of 48% was noted for chromosome 5 (Table 4). When comparing the relative loss of the two arms of a chromosome, there was no significant difference between loss of 18p versus 18q. For the other two chromosomes analyzed, chromosome arms 5q and 17p showed a higher frequency of loss than the corresponding arms 5p and 17q. It should be noted, however, that 5p and 17q were each analyzed with only one DNA probe.

The presence or absence of Type I and II mutations in the tested microsatellite loci were compared to the presence or absence of allelic loss for chromosomes 5, 17, and 18 as determined by Southern blot analysis. A

significant negative correlation was observed between LOH and microsatellite structural changes for each of the three markers evaluated (Table 5). That is, tumors demonstrating genetic instability had infrequent loss of heterozygosity. For example, 14 (22%) tumors did not show LOH for any of the three chromosomes tested. Of these 14 tumors, 12 (80%) exhibited instability at the dinucleotide repeat loci. These results provide further evidence that a genetic mechanism distinct from disruption of known tumor suppressor genes is operative in tumors displaying Type I and Type II mutations.

10

## EXAMPLE 3

Genomic Instability and Anatomical Site of Tumor

Tumors and corresponding normal tissues were surgically removed from patients and the anatomical site of origin of each tumor was recorded. Tumor and normal tissue DNA was extracted and analyzed for Type I or Type II mutations following the method of Example 1 above. Table 6 demonstrates the association of Type I and Type II mutations with anatomical site of the tumor. Thirteen (87%) of the 15 tumors having multiple mutations (i.e. a tumor demonstrating an alteration at more than one microsatellite locus) arose in the proximal colon. Not all of the proximal tumors demonstrated the presence of this type of alteration. That is, Type I or Type II mutational genomic instability was detected in 13 (34%) of the 38 proximal tumors analyzed. Nevertheless, a clear association of multiple microsatellite structural changes and tumor origin in the proximal colon is demonstrated by these results.

25

## EXAMPLE 4

Clinical Significance of Genomic Instability

The clinical significance of the above-described Type I mutation-related genomic instability was examined. Tumor patients having tumors carrying Type I mutations were compared to patients having colorectal tumors without Type I mutations. All tumors were staged by the Astler-Coller modification of the Dukes' staging system. Astler and Collier, The Prognostic Significance of Direct Extension of Carcinoma of the Colon and Rectum. Ann.

30



Surg. 139: 846-51 (1954). Statistical analyses were performed with the Statistical Analysis System (SAS®). Overall survival was defined as the time from surgery until the date of death, while tumor recurrence was defined as the time at which recurrence was documented. Both of these endpoints were censored for patients dying of causes other than metastatic colon cancer. Distributions of survival times were compared using the log-rank statistic, Mantel and Haenszel, Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease, J. Natl. Cancer Inst. 22: 719-48 (1959), and survival distribution curves were estimated by the method of Kaplan and Meier, Nonparametric Estimation from Incomplete Observations, J. Am. Stat. Assoc. 58: 457-81 (1958). The reported p-values are two-sided, and statistical significance was defined at a p-value of less than 0.05.

In this analysis, 86 patients having tumors of Dukes stages A-D were analyzed, and no tumor recurrences nor cause-specific deaths were noted in the 12 patients having Type I mutations. In contrast, approximately 30% of patients without Type I mutations experienced tumor recurrence, and approximately 42% of patients without Type I mutations failed to survive. These data provide a statistically significant association of Type I mutation with decreased disease recurrence ( $p = 0.03$ ) and with increased survival ( $p = 0.02$ ). Sixty patients having tumors classified specifically as Dukes stage B or C were analyzed in the same way. This analysis did not yield a statistically significant association due to the smaller sample size ( $p = 0.07$  for association with decreased disease recurrence and  $p = 0.15$  for association with increased survival). However, the same pattern was seen as with the Dukes A-D patients: no tumor recurrences nor cause-specific deaths in the 11 patients having Type I mutations, while over 23% of the remaining patients experienced tumor recurrence or failed to survive.

#### EXAMPLE 5

##### Analysis of Tri-nucleotide Repeat Locus

DNA was extracted from tumors and normal tissue and standard polymerase chain reactions were carried out using the methods discussed in Example 1 above. A primer pair that binds to the nucleotide sequence flanking a

specific repeat of the form (AGC)<sub>n</sub> at a defined location within the human TATA-binding protein sequence was selected on the basis of a search of a commercial nucleic acid sequence data base. The DNA sequences of these polymerase chain reaction primers are as follows:

5                    5' - ACTGACCCCACAGCCTATTC - 3'  
                    5' - CAAGGGTGCAGTTGTGAGAG - 3'

These primers are sequences numbered 9 and 10 in the DNA Sequence Listing.

As seen in Figure 2, some of the tumors exhibited a structural change in the tested trinucleotide repeat locus when compared to the  
10                    corresponding normal tissues. This can be seen in Figure 2, for example, by comparing tumor sample number 1556 to its corresponding normal tissue sample number 1555. Some tumors carrying one or more di-nucleotide repeat changes failed to show an alteration within the tested tri-nucleotide repeat locus. Nevertheless, only those samples that demonstrated variation in one or more of  
15                    the above-described di-nucleotide ((CA)<sub>n</sub>) repeats also exhibited a structural alteration at the tested tri-nucleotide repeat locus.

Such alterations within a dinucleotide repeat region of the human TATA-binding protein could affect protein function.

20

#### EXAMPLE 6

##### DNA Sequence Analysis of Di-Nucleotide Repeats

DNA sequencing reactions were performed on amplified DNA fragments representing selected di-nucleotide repeat loci in order to confirm the presence of specific structural alterations within each locus. DNA was extracted  
25                    from tumors and normal tissues. Standard polymerase chain reactions using primer pairs for Mfd 41 and Mfd 27 (Table 1) were carried out using the methods discussed in Example 1 above. The amplified DNA's were sequenced using the Sanger method of dideoxy nucleotide sequencing, with the same primers as used for the PCR amplifications. The sequencing protocols were essentially those  
30                    described in the Promega Technical Manual for its *fmol*<sup>™</sup> DNA Sequencing System using Taq DNA polymerase, except that 7-deaza GTP was used in the first PCR reaction. See J. Ellingboe, U.B. Gyllensten, eds., The PCR Technique: DNA

Sequencing (Eaton Publishing, 1992). The labelled DNA was electrophoresed in denaturing polyacrylamide gels. Figures 3 and 4 depict autoradiographs of the sequencing gels.

Figure 3 depicts products of the sequencing reactions from tumor and corresponding tissue DNA of two patients. Prior to sequencing, the DNAs were amplified using the Mfd 41 primer pair. Samples numbered 1555 and 1556 are DNA from tissue and tumor from one patient and samples numbered 1966 and 1967 are tissue and tumor DNA from another patient. When the lanes containing DNA from sample number 1555 (tissue) are compared to those lanes containing DNA from sample number 1556 (tumor), it is apparent that DNA from sample number 1556 is of a higher molecular weight than DNA from sample number 1555. It is also evident from the autoradiograph of the sequencing gel that the amplified DNA, as expected, contains a region of CA-repeats.

Figure 4 depicts products of the sequencing reactions from tumor and corresponding tissue DNA of the same two patients as shown in Figure 3, but with analysis of a different di-nucleotide repeat locus (Mfd 27). The template DNAs were samples numbered 1555, 1556, 1966 and 1967 as discussed above. The sequencing reactions were performed as described above. At this locus, both tumor DNAs varied from the tissue DNAs in the number of di-nucleotide repeats as evidenced by the increased length of the DNA in lanes containing the tumor DNAs. As in Figure 3, Figure 4 also shows a CA-rich region.

## EXAMPLE 7

### Studies on MTS Tumors

The tumors evaluated in Examples 1-6 above were colorectal cancers. The tumors evaluated in this Example represented a variety of neoplasms from MTS patients. Cutaneous biopsy specimens showing sebaceous adenoma, sebaceous epithelioma, or sebaceous carcinoma were identified from a review of all sebaceous neoplasms in the pathology archives of Mayo Clinic from 1923-1983. Finan, M.C., and Connolly, S.M, Sebaceous Gland Tumors and

Systemic Disease: A Clinicopathologic Analysis, Medicine 63:232-242, 1984. Fifty-nine patients had at least one sebaceous tumor of the skin removed. Of these 59 patients, 25 also had at least one internal malignancy and thus were diagnosed as having MTS. Within this group, a total of 49 internal malignancies were identified, 25 of which were colonic, 9 urogenital, 5 hematologic, 4 breast, and one each for pancreas, jejunum, ovary, squamous cell carcinoma of inner ear, paroid (basal cell adenoma), and squamous cell carcinoma of the vocal cord. Additionally, colonic polyps were observed in 12 and a family history of carcinoma was detected in 18. Tumors were staged using the Astler-Coller modification of the Dukes staging system. Astler, V.B., and Collier, F.A., The Prognostic Significance of Direct Extension of Carcinoma of the Colon and Rectum, Ann. Surg. 139:846-851, 1954.

The acquisition of tissue and the review of patient histories were approved by the Mayo Institutional Review Board. As many blocks of paraffin-embedded tumors as possible were obtained from the MTS patients identified by Finan and Connolly. Finan, M.C., and Connolly, S.M., Sebaceous Gland Tumors and Systemic Disease: A Clinicopathologic Analysis, Medicine 63:232-242, 1984. For each block, 5  $\mu$ m thick sections were cut, placed on a glass microscope slide, and standard hematoxylin and eosin staining performed. Areas of the block selected for extraction (normal tissue or tumor tissue containing greater than 40% tumor cells) were carefully etched using a 19 gauge needle. Ten  $\mu$ m thick sections were then cut until approximately 1 cm<sup>2</sup> of tissue was collected. The tissue sections were placed into a microfuge tube and deparaffinized by adding 1 ml of xylol for 10 min. pelleted by centrifugation at 8,000 rpm for 10 min, and the xylol removed with a glass pasteur pipette.

The isolated tissue was washed twice with 1 ml of 95% ethanol and dried under vacuum. Two hundred  $\mu$ l of digestion buffer (10 mM TRIS-HCl [pH 8], 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, and 0.5% Tween 20) and 20  $\mu$ l of glass beads (controlled pore glyceryl glass, Sigma, St. Louis, MO) suspended in H<sub>2</sub>O were added to each tube. The tubes were sonicated for 10 min at 45°C in a Branson 2210 sonicator. Twenty  $\mu$ l of 20 mg Proteinase K/ml H<sub>2</sub>O (Sigma, St. Louis, MO) was added to each tube and incubated overnight at 55°C with shaking. Following

this initial incubation, an additional 10  $\mu$ l of Proteinase K was added to each tube and the tubes incubated for an additional 4 h at 55°C. Two phenol/chloroform and one chloroform extractions were performed on each tube and an aliquot of the extracted aqueous layer was used as template for the polymerase chain  
5 reaction (PCR).

The extracted DNA was examined for genetic alterations at four separate microsatellites localized to chromosome arms 5q (APC), 15q (635/636), 17p (P53) and 18q (Mfd 26). The following primers were utilized for the PCR reaction: SEQ. ID. NO. 11, APC-F (ACTCACTCTAGTGATAAATCG) and SEQ. ID.  
10 NO. 12, APC-R (AGCAGATAAGACAGTATTACTAGTT), allele sizes 96-122 bp; SEQ. ID. NO. 13, 635/636-F (TTGACCTGAATGCACTGTGA) and SEQ. ID. NO. 14, 635/636-R (TTCCATACCTGGGAACGAGT), allele sizes 68-96 bp; SEQ. ID. NO. 15, P53-F (AGGGATACTATTCAGCCCCGAGGTG) and SEQ. ID. NO. 16, P53-R (ACTGCCACTCCTTGCCCCATTC), allele sizes 103-135 bp; and SEQ. ID. NO. 17,  
15 Mfd 26-F (CAGAAAATTCTCTCTGGCTA) and SEQ. ID. NO. 18, Mfd 26-R (CTCATGTTCCTGGCAAGAAT), allele sizes 103-109 bp. PCR and gel electrophoresis was performed essentially as described in Example 1 above.

Age of onset was compared by the exact Wilcoxon rank sum statistic. Conover, W.J., Practical Nonparametric Statistics, 2nd ed. (John Wiley,  
20 New York, 1980) p. 216. Survival distributions were estimated by the Kaplan and Meier method, Kaplan, E.L. and Meier, P., Nonparametric Estimation from Incomplete Observations, J. Am. Statist. Assoc. 53:457-481, 1958, and survival comparison made by the log rank statistic. Mantel, N., Evaluation of Survival Data and Two New Rank Order Statistics Arising in Its Consideration, Cancer  
25 Chemother. Rep. 50:163-170, 1966.

Thirteen MTS patients were assessed for MIN in at least one sebaceous tumor or one visceral malignancy. The clinical and pathological findings in these 13 patients are summarized in Table 7. Those tumors with PCR data are indicated with an asterisk and a summary of these analyses is shown in  
30 Table 8. Six patients had tumors which exhibited MIN in at least 3 of the 4 loci tested (Group I). Examples of MIN in a variety of tumors are shown in Figure 5. One patient (case 21) exhibited MIN at only 1 of the 4 loci in a single tumor. The

remaining 6 patients failed to demonstrate MIN in any of the tumors examined (Group II). The age of onset for the first visceral malignancy, survival after the onset of this malignancy, and presence or absence of family history of cancer for the 13 patients are summarized in Table 9.

5           Patients with tumors that demonstrated MIN had an average of 4.8 sebaceous tumors, 19 skin tumors of other types and 3.5 visceral malignancies, while patients with tumors lacking MIN had an average of 1.0 sebaceous tumors, 1.3 skin tumors of other types and 1.3 visceral malignancies. The average age of diagnosis for the first sebaceous tumor was 62 years (range: 55-73 y) in Group I  
10   and 70 years (range: 57-85 y) in Group II. In all Group I patients, the first sebaceous tumor occurred after the development of the first visceral malignancy (average of 22 y with a range of 2-37 y). On the other hand, the average age at which the first sebaceous tumor was diagnosed in a Group II patient was the same as the age of diagnosis for the first visceral malignancy (range: 11 y before  
15   to 11 y after). No difference was observed in the number of colonic polyps detected in patients with or without MIN.

Overall, the findings suggested that the underlying molecular basis of MTS is likely to be heterogeneous in nature, with the presence of at least two subgroups of patients: those with (Group I) and those without (Group II)  
20   widespread MIN in their tumors. A comparison of the clinical and pathologic features of these two groups of patients revealed several interesting differences.

First, Group I patients had an earlier age of onset of their first visceral malignancy (average age of 40 y, range 23-53 y) compared to Group II patients (average age 70 y, range 62-79 y), a difference that was highly significant  
25   ( $p=0.001$ ). Second, survival time, with death from any cause, was significantly shorter ( $p=0.04$ ) for patients without MIN (median survival of 11 y) compared to patients with MIN (median survival of 32 y). These results were based on the patient population which had 4 deaths/ 6 patients in Group I and 5 deaths/ 7 patients in Group II. One patient in each group died without evidence of cancer  
30   (one with Alzheimer's disease and one from heart disease). Third, all Group I patients had at least one colorectal cancer as one of their visceral malignancies while this was not necessarily the case for Group II patients. Among the 4

patients in Group 2 not having colon cancer, two had breast cancer, one had small lymphocytic lymphoma and another a bladder cancer. Fourth, all Group I patients had a family history of cancer while this was true for only four of the seven patients in Group II. Finally, the development of numerous additional skin and visceral malignancies was common in Group I while uncommon in Group II.

The foregoing detailed description has been provided for a better understanding of the invention only and no unnecessary limitation should be understood therefrom as some modifications will be apparent to those skilled in the art without deviating from the spirit and scope of the appended claims.

Table 1: Primers for Amplification of Di-nucleotide Repeats

<u>Marker</u>	<u>Locus</u>	<u>Map Position</u>	<u>Seq. No.</u>	<u>Primer Sequence</u>
Mfd 27	D5S107	5q11.2-q13.3	1	GATCCACTTTAACCCTAAATAC
			2	GGCATCAACTTGAACAGCAT
Mfd 41	D17S261	17p12-p11.1	3	CAGGTTCTGTCATAGGACTA
			4	TTCTGGAAACCTACTCCTGA
Mfd 26	D18S34	18q	5	CAGAAAATTCTCTCTGGCTA
			6	CTCATGTTCTCTGGCAAGAAT
635/636	cardia muscle actin gene	15q11-qter	7	TTGACCTGAATGCACTGTGA
			8	TTCCATACCTGGGAACGAGT

Mfd 26, 27, and 41 sequence obtained from James Weber. 635/636 sequence obtained from M. Litt and J.A. Luty, A. Hypervariable Microsatellite Revealed by *in vitro* Amplification of a Dinucleotide Repeat Within the Cardiac Muscle Actin Gene, Am. J. Human Genet. 44:397 (1989).



Table 2 -- Tumors demonstrating genomic instability.

Type of Alteration	<u>Number of Loci Involved/Tumor</u>			
	1	2	3	4
I alone	0	1	3 (2)	4 (3)
II alone	10	2	0	0
I and II	0	0	0	5 (4)

( ) = Synchronous tumors omitted

Table 3: Chromosomal Location of DNA Probes Utilized

<u>Probe</u>	<u>Locus Symbol</u>	<u>Location</u>	<u>Restriction Enzyme</u>	<u>Reference</u>
pJO71H-A	D5S20	5p13	EcoRI	1
p105-153A	D5S39	5q11.2-q13.3	MspI	1
$\pi$ 227	D5S37	5q21	PstI	1
C11p11	D5S71	5q14-q21	TaqI	1
M4	D5S6	5q11.2-q13.3	BamHI	1
J0205H-C	D5S22	5q34-qter	MspI	1
pYNZ22	D17S30	17p13.3	BamHI	2
pYNH37.3	D17S28	17p13.3	TaqI	1
pTHH59	D17S4	17q23-q25.3	TaqI	1
L2.7	D18S6	18p11	PstI	1
pHH64	TTR	18q11.2-q12.1	MspI	1
DCC1.9	DCC	18q21.3	EcoRI	4
p15-65	D18S8	18q21.3	MspI	3
pERT25	D18S11	18q23	PstI	1

1. Cytogenetics and Cell Genetics 58: 1-2200 (1991).
2. Nakamura Y: Nuc Acids Res 16:4707 (1987).
3. Marthens F, et al. Nuc. Acids Res 15:1348 (1987).
4. Vogelstein, unpublished results.

Table 4: Frequency of Allelic Loss on Chromosomes 5, 17 and 18

Chromosome	Loss		No Loss (No.)	NI (No.)
	No.	% *		
5p	9	20	36	46
5q	40	46	47	4
5	42	48	46	3
17p	59	69	26	6
17q	29	45	36	26
17	61	69	28	2
18p	24	69	11	56
18q	59	69	27	5
18	62	70	27	2

NI = non-informative

\* percentage refers to the number of tumors demonstrating loss of heterozygosity divided by the total number of tumors that were informative for the DNA probes utilized.

Table 5: Association of DNA Alterations with Loss of Heterozygosity

		NL	L		
		5q			
MFD 27	Neg	34 (33)	37 (36)	p=0.025	
	I	11 (8)	1		
	II	1	2		
		17p			
MFD 41	Neg	17 (16)	56 (55)	p<0.0005	
	I	7 (6)	1 (0)		
	II	2	0		
		18q			
MFD 26	Neg	15 (14)	53 (51)	p<0.0005	
	I	10 (8)	0		
	II	1	4		

NL = no loss

L = loss

( ) synchronous tumors omitted

Table 6 -- Association of DNA alterations with anatomical site.

	C	AC	HF	TC	SF	DC	S	R	P	D	
MFD27	Neg I II	4 4 (3) 1	2 1 0	5 1 (0) 0	1 0 0	2 0 0	20 (19) 0 0	26 (25) 1 2	26 11 (8) 1	49 (47) 1 2	p=0.003
635	Neg I II	3 4 2 (1)	2 0 1	5 1 (0) 0	1 0 0	2 0 0	19 (18) 0 1	23 (22) 1 5	25 8 (6) 5 (4)	45 (43) 1 6	p=0.018
MFD41	Neg I II	4 5 (4) 0	2 0 1	6 (5) 0 0	1 0 0	2 0 0	20 (19) 0 0	28 0 0	28 (27) 8 (6) 2	51 (50) 0 0	p=0.003
MFD26	Neg I II	4 4 (3) 1	1 1 0	5 1 (0) 0	1 0 0	2 0 0	18 (17) 0 2	27 (26) 1 1	25 10 (7) 2	48 (46) 1 3	p=0.006
Tumors demonstrating mutations at $\leq 1$ or $\geq 2$ loci		3 6 (5)	1 1	5 1 (0)	1 0	2 0	20 (19) 0	26 2	24 13 (10)	49 (48) 2	p=0.003
	C AC HF TC SF	caecum ascending colon hepatic flexure transverse colon splenic flexure					DC S R P D		( )	with synchronous tumors omitted	

35

Table 7. Clinicopathologic findings in MTS patients with tumors exhibiting microsatellite instability (Group I: cases 4, 6, 9, 23, 24, and 25) and with tumors not exhibiting microsatellite instability (Group II: cases 3, 5, 11, 17, 19, 21, and 22).

Case #	Number of sebaceous tumor(s)	Other skin tumors	Visceral malignancy			
			Site	Histology	Stage	Age <sup>o</sup>
4	5*	BCC KA* SK (12) VK (14)	Cecum*	ACA	B2	53
6	7*	AK (>50) IFK KA SCC (2) SK (3) VK	Ascending colon Rectosigmoid colon* Bladder Inner ear Kidney* Bladder Kidney Liver	? (outside institution) ACA TCC SCC RCC TCC RCC Metastatic ACA	? B2	42 47 49 72 73 74 74 74
9	5*	BCC (2) KA (2) SCC (4) SK VK (3)	Transverse colon Lymph node* Ascending colon* Lymph node	ACA CLL ACA CLL	B2 B2	38 62 67 69
23	6*	AK (2) SCC VK	Rectum* Ascending colon Transverse colon Kidney (renal pelvis)* Prostate	ACA ACA ACA TCC ACA	B1 B2 B2	42 64 65 70 73
24	5*	AK (2) EC (5) IFK	Cecum Rectum Liver Ovary Cecum* Rectum*	ACA ACA Metastatic ACA Adenoacanthoma ACA ACA	C2 C2 D B2 B1	23 29 30 33 51 64
25	1*	AK (3)	Descending colon Ascending colon* Retropertoneum	ACA ACA Metastatic ACA	C2 B1 D	40 52 62
3	1*	BCC (2)* Melanoma SK	Bladder	TCC		79
5	1	....	Colon (splenic flexure) Colon (hepatic flexure)* Peritoneum	ACA ACA Metastatic ACA	C2 C2 D	69 79 80
11	1	....	Cecum* Lymph node (groin)	ACA NHL	B2	79 93
17	1	EC SK	Stomach* Cervix Stomach*	NHL In situ SCC NHL		62 64 69
19	1*	EC	Breast*	ACA		68
21	1	AK (2)	Ascending colon* Jejunum* Peritoneum	ACA Metastatic ACA Metastatic ACA	A D D	68 68 68
22	1*	....	Breast*	Invasive ACA		74

## KEY TO ABBREVIATIONS AND SYMBOLS

ACA-adenocarcinoma  
AK-actinic keratosis  
BCC-basal cell carcinoma  
CLL-chronic lymphocytic leukemia  
EC-epidermoid cyst  
IFK-inverted follicular keratosis  
KA-keratoacanthoma  
NHL-non-Hodgkin's lymphoma

RCC-renal cell carcinoma  
SCC-squamous cell carcinoma  
SK-seborrheic keratosis  
TCC-transitional cell carcinoma  
VK-verrucal keratosis  
\*PCR data obtained  
†Stage of c. colorectal cancer  
°Age at diagnosis of visceral malignancy

Table 8. Results of assays for microsatellite instability in the various tumors.

## PATIENTS WITH MICROSATELLITE INSTABILITY †

Case #	Tumor site (histology)*	# of loci demonstrating microsatellite instability
4	Cecum (ACA)	4/4
	Skin (SE)	4/4
	Skin (KA)	4/4
6	Rectosigmoid colon (ACA)	3/4
	Kidney (RCC)	0/4
	Skin (SA)	4/4
	Skin (SA)	3/4
9	Skin (SA)	4/4
	Skin (SA)	3/4
	Ascending Colon (ACA)	4/4
	Skin (SC)	4/4
	Skin (SC)	4/4
	Lymph node (CLL)	0/4
23	Rectum (ACA)	4/4
	Skin (SA)	4/4
	Skin (SA)	4/4
	Renal pelvis (TCC)	3/4
	Skin (SC)	4/4
	Prostate (ACA)	3/4
24	Cecum (ACA)	3/4
	Skin (SA)	4/4
	Rectum (ACA)	3/4
25	Ascending colon (ACA)	4/4
	Skin (SA)	3/4

## PATIENTS WITHOUT MICROSATELLITE INSTABILITY

Case #	Tumor site (histology)	# of loci demonstrating microsatellite instability
3	Skin (BCC)	0/4
	Skin (SC)	0/4
5	Hepatic flexure (ACA)	0/4
11	Cecum (ACA)	0/4
17	Stomach (NHL)	0/4
	Stomach (NHL)	0/4
19	Eyelid (Meibomian gland adenoma)	0/4
	Breast (ACA)	0/4
21	Ascending colon (ACA)	0/4
	Jejunum (metastatic ACA)	1/4
22	Eyelid (Meibomian gland carcinoma)	0/4
	Breast (ACA)	0/4

## KEY TO ABBREVIATIONS AND SYMBOLS

\*Tumors are ordered according to sequence of occurrence in patient.

†Instability at 3 of 4 loci

SA-sebaceous adenoma

SE-sebaceous epithelioma

SC-sebaceous carcinoma

See Table 1 for abbreviations of remaining tumor types.

Table 9. Summary of clinical features of patients with or without microsatellite instability.

## PATIENTS WITH MICROSATELLITE INSTABILITY†

Case #	Age of onset of first visceral malignancy (tumor type)	Survival following onset of first visceral malignancy (y)	Family history of cancer	Comments
4	53 (Cecum)	24	Yes	Alive, no evidence of malignancy at age 78.
6	42 (Ascending colon)	32	Yes	Death from renal cell carcinoma* at age 74.
9	38 (Transverse colon)	32	Yes	Death from CLL* at age 69.
23	42 (Rectum)	35	Yes	Death from Alzheimer's disease at age 76. No malignancy at death.
24	23 (Cecum)	37	Yes	Alive, no evidence of malignancy at age 70.
25	40 (Descending colon)	23	Yes	Death from suicide at age 63. Metastatic colon cancer at death.

## PATIENTS WITHOUT MICROSATELLITE INSTABILITY

Case #	Age of onset of first visceral malignancy (tumor type)	Survival following onset of first visceral malignancy (y)	Family history of cancer	Comments
3	79 (Bladder)	9	Yes	Death from metastatic melanoma at age 89.
5	69 (Splenic flexure)	11	No	Death from metastatic colon cancer at age 81.
11	70 (Cecum)	25	Yes	Lost to follow up at age 95. High stage lymphoma at that time.
17	62 (Lymphoma)	23	No	Alive, no evidence of malignancy at age 85.
19	68 (Breast)	11	No	Death at age 79 from "heart disease".
21	68 (Ascending colon)	1	Yes	Death from metastatic colon cancer at age 69.
22	74 (Breast)	1	Yes	Death from metastatic Melibonolan gland carcinoma at age 75.

†Instability in at least 3 of the 4 loci

\* Only tumor not exhibiting microsatellite instability in this patient.



SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- 5 (i) APPLICANT: Stephen N. Thibodeau  
Gary D. Bren
- (ii) TITLE OF INVENTION: TUMOR-SPECIFIC GENOMIC INSTABILITY  
AS A PROGNOSTIC INDICATOR
- 10 (iii) NUMBER OF SEQUENCES: 18
- (iv) CORRESPONDENCE ADDRESS:
- 15 (A) ADDRESSEE: Patterson & Keough, P.A.  
(B) STREET: 1200 Rand Tower  
527 Marquette Avenue South
- 20 (C) CITY: Minneapolis  
(D) STATE: Minnesota  
(E) COUNTRY: USA
- 25 (F) ZIP: 55402
- (v) COMPUTER READABLE FORM:
- 30 (A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: Apple Macintosh  
(C) OPERATING SYSTEM: Apple Macintosh System 7.0.1
- 35 (D) SOFTWARE: WordPerfect 2.1.4 for the Macintosh
- (vi) CURRENT APPLICATION DATA:
- 40 (A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:
- 45 (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Mark S. Ellinger, Esq.
- 50 (B) REGISTRATION NUMBER: 34,812  
(C) REFERENCE/DOCKET NUMBER: 1144.01-WO-01
- 55

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 612/349-5743

5

(B) TELEFAX: 612/349-9266

## 41

## (2) INFORMATION FOR SEQ ID NO:1

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 21  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
10 (D) TOPOLOGY: linear
- (viii) POSITION IN GENOME:  
15 (A) CHROMOSOME/SEGMENT: 5q11.2-q13.3  
(B) MAP POSITION: D5S107

20 (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:1:

GATCCACTTT AACCCAAATA C 21

25

## (2) INFORMATION FOR SEQ ID NO:2

## (i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 20  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
35 (D) TOPOLOGY: linear
- (viii) POSITION IN GENOME:  
40 (A) CHROMOSOME/SEGMENT: 5q11.2-q13.3  
(B) MAP POSITION: D5S107

45

## (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:2:

GGCATCAACT TGAACAGCAT 20

42

## (2) INFORMATION FOR SEQ ID NO:3

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 20  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
10 (D) TOPOLOGY: linear

## (viii) POSITION IN GENOME:

- 15 (A) CHROMOSOME/SEGMENT: 17p12-p11.1  
(B) MAP POSITION: D17S261

20 (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:3:

CAGGTTCTGT CATAGGACTA 20

25

## (2) INFORMATION FOR SEQ ID NO:4

## (i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 20  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
35 (D) TOPOLOGY: linear

40 (viii) POSITION IN GENOME:

- (A) CHROMOSOME/SEGMENT: 17p12-p11.1  
(B) MAP POSITION: D17S261

45

## (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTCTGGAAAC CTACTCCTGA 20

43

## (2) INFORMATION FOR SEQ ID NO:5

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 20  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
10 (D) TOPOLOGY: linear

## (viii) POSITION IN GENOME:

- 15 (A) CHROMOSOME/SEGMENT: 18q  
(B) MAP POSITION: D18S34

20

## (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:5:

CAGAAAATTC TCTCTGGCTA 20

25

## (2) INFORMATION FOR SEQ ID NO:6

## (i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 20  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
35 (D) TOPOLOGY: linear

40

## (viii) POSITION IN GENOME:

- (A) CHROMOSOME/SEGMENT: 18q  
(B) MAP POSITION: D18S34

45

## (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:6:

CTCATGTTCC TGGCAAGAAT 20

44

## (2) INFORMATION FOR SEQ ID NO:7

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 20  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
10 (D) TOPOLOGY: linear

## (viii) POSITION IN GENOME:

- 15 (A) CHROMOSOME/SEGMENT: 15q11-qter

## (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:7:

20 TTGACCTGAA TGCACTGTCA 20

## 25 (2) INFORMATION FOR SEQ ID NO:8

## (i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 20  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
35 (D) TOPOLOGY: linear

## (viii) POSITION IN GENOME:

- 40 (A) CHROMOSOME/SEGMENT: 15q11-qter

## (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:8:

45 TTCCATACCT GGCAACGAGT 20

45

## (2) INFORMATION FOR SEQ ID NO:9

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 20  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
10 (D) TOPOLOGY: linear

## (x) PUBLICATION INFORMATION:

- 15 (A) AUTHORS: Peterson, M.G.  
Tanese, N.  
Pugh, B.F.  
Tjian, R.  
20 (B) TITLE: Functional domains and upstream  
activation properties of cloned human TATA  
binding protein  
(C) JOURNAL: Science  
25 (D) VOLUME: 248  
(F) PAGES: 1625-1630  
30 (G) DATE: 1990  
(K) RELEVANT RESIDUES: 352 - 371

## 35 (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:9:

ACTGACCCCA CAGCCTATTC 20

40

## (2) INFORMATION FOR SEQ ID NO:10

## (i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 20  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
50 (D) TOPOLOGY: linear

## (x) PUBLICATION INFORMATION:

- 5 (A) AUTHORS: Peterson, M.G.  
Tanese, N.  
Pugh, B.F.  
Tjian, R.
- 10 (B) TITLE: Functional domains and upstream  
activation properties of cloned human TATA  
binding protein
- (C) JOURNAL: Science
- 15 (D) VOLUME: 248
- (F) PAGES: 1625-1630
- (G) DATE: 1990
- 20 (K) RELEVANT RESIDUES: 618 - 637

## (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:10:

25 CAAGGGTGCA GTTGTGAGAG 20

## (2) INFORMATION FOR SEQ ID NO:11

- 30 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 21
- 35 (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- 40 (D) TOPOLOGY: linear

## (viii) POSITION IN GENOME:

- 45 (A) CHROMOSOME/SEGMENT: 5q

## (x) PUBLICATION INFORMATION:

- (A) AUTHORS: Spiro, L. et al.
- 50 (B) TITLE: A CA repeat 30-70 Kb downstream from the  
adenomatous polyposis coli (APC) gene
- (C) JOURNAL: Nucleic Acids Res.
- 55 (D) VOLUME: 19



47

(F) PAGES: 6348 et. seq.

(G) DATE: 1991

5

(xi) A SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACTCACTCTA GTGATAAATC G 21

48

## (2) INFORMATION FOR SEQ ID NO:12

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 25  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
10 (D) TOPOLOGY: linear

## (viii) POSITION IN GENOME:

- 15 (A) CHROMOSOME/SEGMENT: 5q

## (x) PUBLICATION INFORMATION:

- 20 (A) AUTHORS: Spiro, L. et al.  
(B) TITLE: A CA repeat 30-70 Kb downstream from the  
adenomatous polyposis coli (APC) gene  
25 (C) JOURNAL: Nucleic Acids Res.  
(D) VOLUME: 19  
(F) PAGES: 6348 et. seq.  
30 (G) DATE: 1991

## (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:12:

35 AGCAGATAAG ACAGTATTAC TAGTT 25

## 40 (2) INFORMATION FOR SEQ ID NO:13

## (i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 20  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
50 (D) TOPOLOGY: linear

## (viii) POSITION IN GENOME:

- 55 (A) CHROMOSOME/SEGMENT: 15q

## (x) PUBLICATION INFORMATION:

- 5 (A) AUTHORS: Thibodeau, S.N. et al.
- (B) TITLE: Microsatellite instability in cancer of  
the proximal colon
- 10 (C) JOURNAL: Science
- (D) VOLUME: 260
- (F) PAGES: 816-819
- 15 (G) DATE: 1993

## (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:13:

20 TTGACCTGAA TGCACTGTGA 20

50

## (2) INFORMATION FOR SEQ ID NO:14

## (i) SEQUENCE CHARACTERISTICS:

5

(A) LENGTH: 20

(B) TYPE: nucleic acid

10

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (viii) POSITION IN GENOME:

15

(A) CHROMOSOME/SEGMENT: 15q

## (x) PUBLICATION INFORMATION:

20

(A) AUTHORS: Thibodeau, S.N. et al.

(B) TITLE: Microsatellite instability in cancer of the proximal colon

25

(C) JOURNAL: Science

(D) VOLUME: 260

30

(F) PAGES: 816-819

(G) DATE: 1993

## (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:14:

35

TTCCATACCT GGGAACGAGT 20

## 40 (2) INFORMATION FOR SEQ ID NO:15

## (i) SEQUENCE CHARACTERISTICS:

45

(A) LENGTH: 24

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

50

(D) TOPOLOGY: linear

## (viii) POSITION IN GENOME:

55

(A) CHROMOSOME/SEGMENT: 17p

51

## (x) PUBLICATION INFORMATION:

- 5 (A) AUTHORS: Jones, M.H., and Nakamura, Y.
- (B) TITLE: Detection of loss of heterozygosity at  
the human TP53 locus using a dinucleotide  
repeat polymorphism
- 10 (C) JOURNAL: Genes Chrom. Cancer
- (D) VOLUME: 5
- (F) PAGES: 89-90
- 15 (G) DATE: 1992

## (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:15:

AGGGATACTA TTCAGCCCCGA GGTG 24

52

## (2) INFORMATION FOR SEQ ID NO:16

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 22
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- 10 (D) TOPOLOGY: linear

## (viii) POSITION IN GENOME:

- 15 (A) CHROMOSOME/SEGMENT: 17p

## (x) PUBLICATION INFORMATION:

- 20 (A) AUTHORS: Jones, M.H., and Nakamura, Y.
- (B) TITLE: Detection of loss of heterozygosity at  
the human *TP53* locus using a dinucleotide  
repeat polymorphism
- 25 (C) JOURNAL: Genes Chrom. Cancer
- (D) VOLUME: 5
- 30 (F) PAGES: 89-90
- (G) DATE: 1992

## (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:16:

ACTGCCACTC CTTGCCCCAT TC 22

40

## (2) INFORMATION FOR SEQ ID NO:17

## (i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 20
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- 50 (D) TOPOLOGY: linear

## (viii) POSITION IN GENOME:

55

53

(A) CHROMOSOME/SEGMENT: 18q

## (x) PUBLICATION INFORMATION:

5 (A) AUTHORS: Thibodeau, S.N. et al.

(B) TITLE: Microsatellite instability in cancer of  
the proximal colon

10 (C) JOURNAL: Science

(D) VOLUME: 260

15 (F) PAGES: 816-819

(G) DATE: 1993

20 (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:17:

CAGAAAATTC TCTCTGGCTA 20

54

## (2) INFORMATION FOR SEQ ID NO:18

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 20  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
10 (D) TOPOLOGY: linear

## (viii) POSITION IN GENOME:

- 15 (A) CHROMOSOME/SEGMENT: 18q

## (x) PUBLICATION INFORMATION:

- 20 (A) AUTHORS: Thibodeau, S.N. et al.  
(B) TITLE: Microsatellite instability in cancer of  
the proximal colon  
25 (C) JOURNAL: Science  
(D) VOLUME: 260  
(F) PAGES: 816-819  
30 (G) DATE: 1993

## (xi) A SEQUENCE DESCRIPTION: SEQ ID NO:18:

35 CTCATGTTCC TGGCAAGAAT 20



We claim:

- 1 1. A method for detecting tumor-specific genomic instability in a  
2 malignant tumor as a prognostic indicator for a patient having said malignant  
3 tumor, comprising:
  - 4 (a) obtaining a sample of said tumor, said tumor sample  
5 having an operable fraction of tumorigenic cells;
  - 6 (b) obtaining a tissue sample from said patient, said tissue  
7 sample being operably free of tumorigenic cells;
  - 8 (c) obtaining genomic DNA from said tumor sample and  
9 from said tissue sample;
  - 10 (d) employing means for comparing the structure of at  
11 least one interspersed repetitive sequence locus in said tumor  
12 sample DNA with the structure of the corresponding at least one  
13 interspersed repetitive sequence locus in said tissue sample DNA to  
14 determine the structural status of said at least one locus in said  
15 tumor sample DNA; and
  - 16 (e) correlating said structural status with the presence or  
17 absence of said genomic instability in said malignant tumor.
- 1 2. The method of claim 1 wherein said interspersed repetitive  
2 sequence is a microsatellite sequence.
- 1 3. The method of claim 2 wherein said microsatellite sequence is a di-  
2 nucleotide repeat sequence.
- 1 4. The method of claim 2 wherein said microsatellite sequence is a tri-  
2 nucleotide repeat sequence.
- 1 5. The method of claim 3 wherein said di-nucleotide repeat sequence  
2 is of the form  $(CA)_n (GT)_n$ .

1 6. The method of claim 1 wherein said malignant tumor is a colorectal  
2 tumor.

1 7. The method of claim 1 wherein said comparison of said at least one  
2 interspersed repetitive sequence locus comprises selecting appropriate primers  
3 for polymerase-based amplification of said at least one locus, amplifying said at  
4 least one locus in said tumor sample DNA and in said tissue sample DNA, and  
5 comparing the apparent fragment sizes of said amplified DNA's.

1 8. The method of claim 7 wherein said at least one locus comprises at  
2 least two unlinked loci.

1 9. The method of claim 7 wherein said polymerase-based  
2 amplification is by polymerase chain reaction.

1 10. A method for evaluating the clinical prognosis of a patient having a  
2 malignant tumor, comprising:

3 (a) obtaining a sample of said tumor, said tumor sample  
4 having an operable fraction of tumorigenic cells;

5 (b) obtaining a tissue sample from said patient, said tissue  
6 sample being operably free of tumorigenic cells;

7 (c) obtaining genomic DNA from said tumor sample and  
8 from said tissue sample;

9 (d) employing means for comparing the structure of at  
10 least one interspersed repetitive sequence locus in said tumor  
11 sample DNA with the structure of the corresponding at least one  
12 interspersed repetitive sequence locus in said tissue sample DNA to  
13 determine the structural status of said at least one locus in said  
14 tumor sample DNA; and

15 (e) correlating said structural status with the clinical  
16 prognosis of said patient.

1 11. The method of claim 10 wherein said interspersed repetitive  
2 sequence is a microsatellite sequence.

1 12. The method of claim 11 wherein said microsatellite sequence is a di-  
2 nucleotide repeat sequence.

1 13. The method of claim 11 wherein said microsatellite sequence is a  
2 tri-nucleotide repeat sequence.

1 14. The method of claim 12 wherein said di-nucleotide repeat sequence  
2 is of the form  $(CA)_n \cdot (GT)_n$ .

1 15. The method of claim 10 wherein said malignant tumor is a  
2 colorectal tumor.

1 16. The method of claim 10 wherein said comparison of said at least  
2 one interspersed repetitive sequence locus comprises selecting appropriate  
3 primers for polymerase-based amplification of said at least one locus, amplifying  
4 said at least one locus in said tumor sample DNA and in said tissue sample  
5 DNA, and comparing the apparent fragment sizes of said amplified DNA's.

1 17. The method of claim 16 wherein said at least one locus comprises at  
2 least two unlinked loci.

1 18. The method of claim 16 wherein said polymerase-based  
2 amplification is by polymerase chain reaction.

1 19. An article of manufacture, comprising packaging material and the  
2 means of claim 1 or claim 10 for comparing the structure of said at least one  
3 interspersed repetitive sequence locus, said means being carried within said  
4 packaging material, wherein said packaging material comprises a label that

5 indicates that said means can be used for assessing tumor-specific genomic  
6 instability in a malignant tumor.

1/4

Fig. 1A

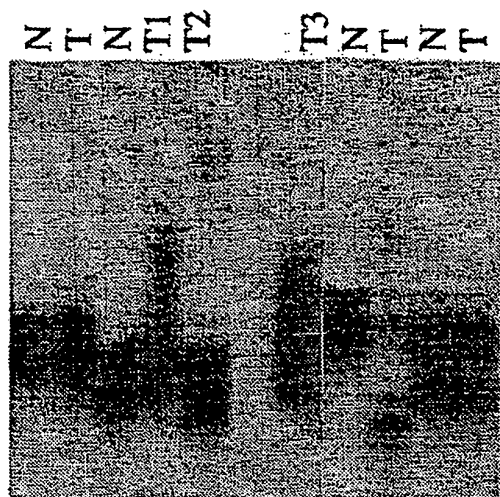


Fig. 1B

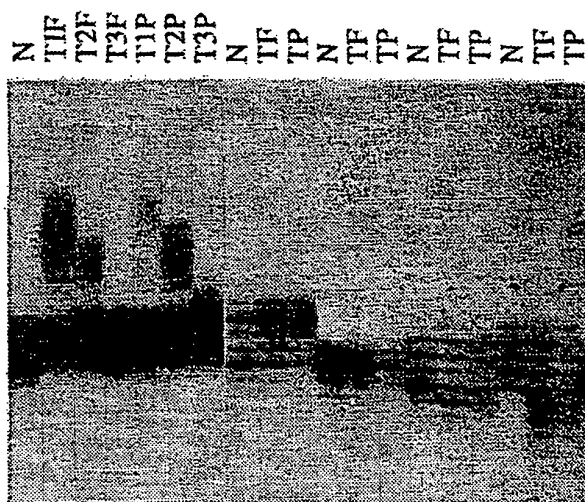
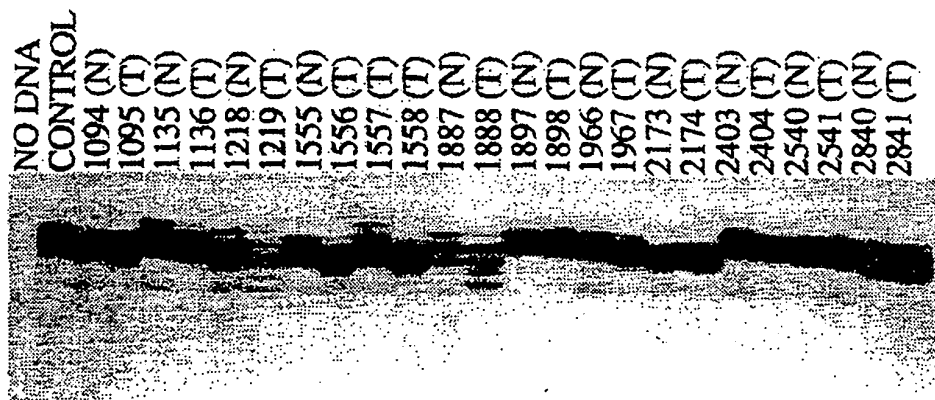


Fig. 2

TFIID (CAG)<sub>n</sub>



2/4  
Fig. 3  
Mfd 41

G				A				T				C			
CONTROL	1555 (N)	1556 (T)	1966 (N)	1967 (T)	CONTROL	1555 (N)	1556 (T)	1966 (N)	1967 (T)	CONTROL	1555 (N)	1556 (T)	1966 (N)	1967 (T)	

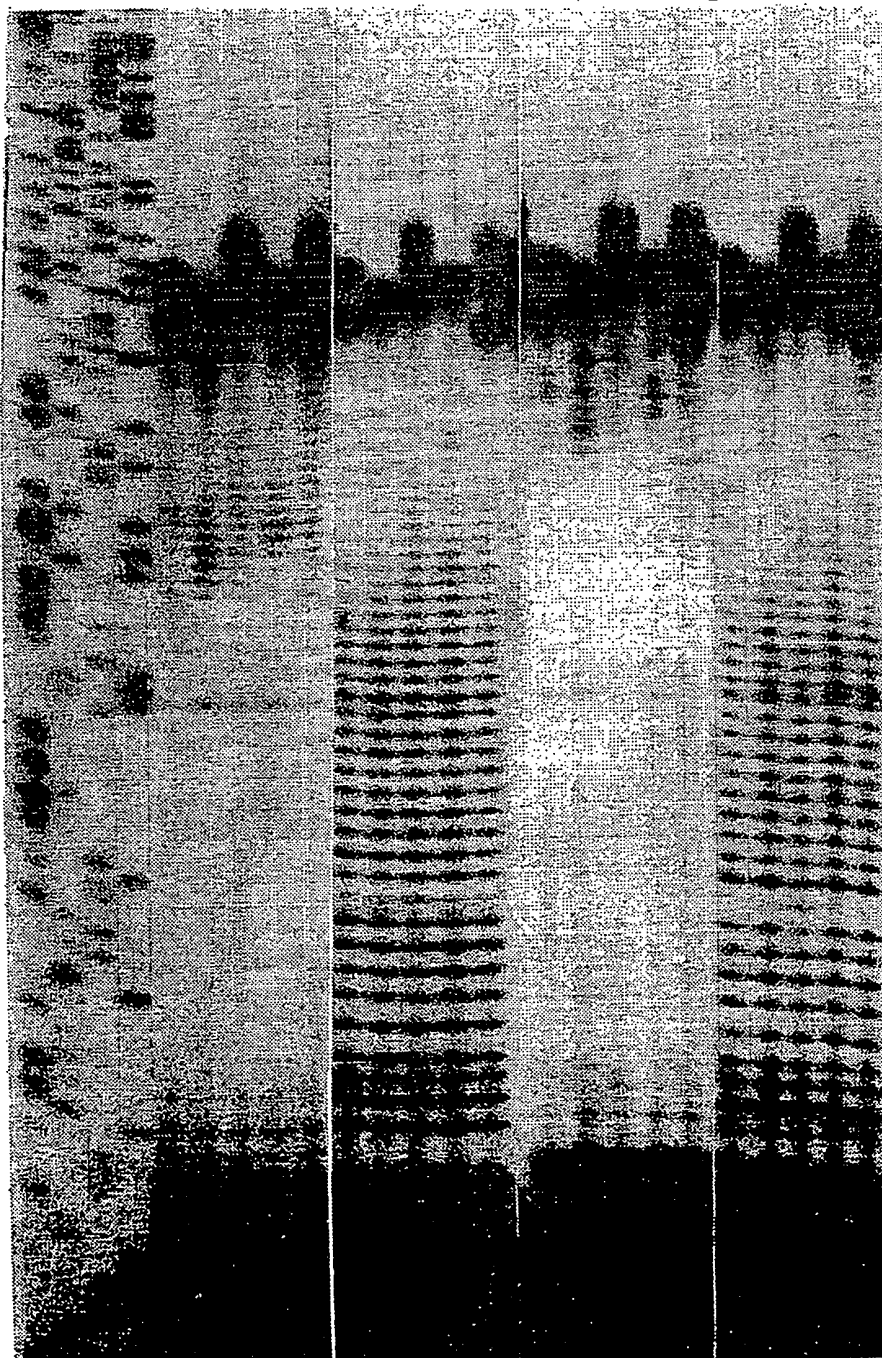


3/4

Fig. 4

Mfd 27

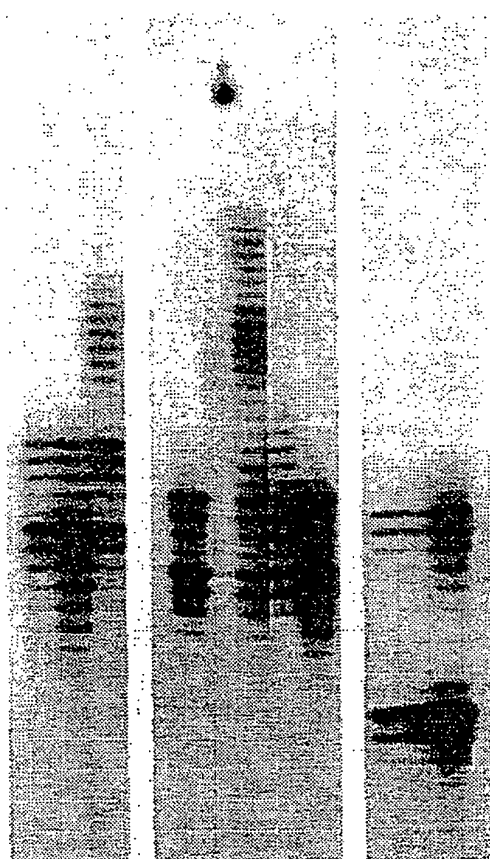
G				A				T				C			
CONTROL	1555 (N)	1556 (T)	1966 (N)	1967 (T)	CONTROL	1555 (N)	1556 (T)	1966 (N)	1967 (T)	CONTROL	1555 (N)	1556 (T)	1966 (N)	1967 (T)	



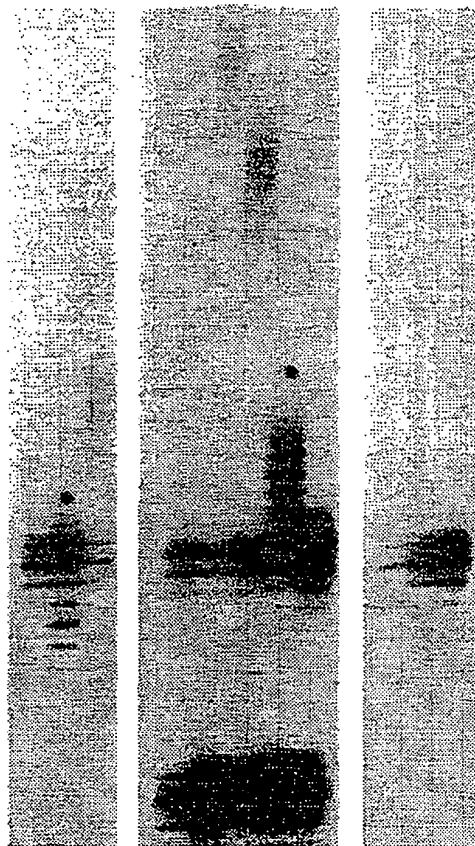
4/4

*Fig. 5A*

Case 4    Case 9    Case 19

N 12  
++N 1 2 3 4  
+++ -N 12  
--*Fig. 5B*

Case 4    Case 9    Case 19

N 12  
++N 1 2 3 4  
+++ -N 12  
--



## INTERNATIONAL SEARCH REPORT

Int. national application N .

PCT/US94/01761

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(5) : C12Q 1/68; C12P 19/34

US CL : 435/6, 91.2

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 91.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

BIOSIS, CURR. BIOTECH ABSTR., DERWENT BIOTECH ABSTR., EMBASE, MEDLINE, WPI, LIFE SCI. ABSTR.,  
CHEM. ABSTR., DISSERTATION ABSTR.**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	CANCER RESEARCH, Volume 53, issued 15 December 1993, Peltomaki et al., "Microsatellite Instability is Associated with Tumors That Characterize the Hereditary Non-Polyposis Colorectal Carcinoma Syndrome", pages 5853-5855, see pages 5853-5854.	1-18
P, X	CANCER RESEARCH, Volume 53, issued 15 December 1993, Lothe et al., "Genomic Instability in Colorectal Cancer: Relationship to Clinicopathological Variables and Family History", pages 5849-5852, see pages 5850-5851.	1-18

☒ Further documents are listed in the continuation of Box C.
 ☐ See patent family annex.

* Special categories of cited documents:	T	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Z"	document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means		
"P" document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 03 JUNE 1994	Date of mailing of the international search report 17 JUN 1994
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer STEPHANIE W. ZITOMER, PH.D. Telephone No. (703) 308-0196

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US94/01761

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CELL, Volume 27, issued December 1981 (Part 2), Perucho et al., "Human-Tumor-Derived Cell Lines Contain Common and Different Transforming Genes", pages 467-476, see pages 468-469, 475, 476 (Note Added in Proof).	1-19
Y	GENE, Volume 108, issued 1991, Bernues et al., "SV40 Recombinants Carrying a d(CT.GA) <sub>22</sub> Sequence Show Increased Genomic Instability", pages 269-274, see pages 269-270, 273.	1-19
Y	AMERICAN JOURNAL OF HUMAN GENETICS, Volume 44, issued 1989, Weber et al., "Abundant Class of Human DNA Polymorphisms Which Can Be Typed Using the Polymerase Chain Reaction", pages 388-396, see pages 388-390, 393-395.	1-19